

NOTICE OF EXPRESS MAILING

Express Mail Mailing Label Number: EE802420245US

Date of Deposit with USPS: January 16, 2002

APPLICATION FOR LETTERS PATENT

for

**SELF-ALIGNED TRENCH MOSFETS AND METHODS FOR
MAKING THE SAME**

Inventors:

Duc Chau
Becky Losee
Bruce Marchant
Dean Probst
Robert Herrick
Jim Murphy

**SELF-ALIGNED TRENCH MOSFETS AND METHODS
FOR MAKING THE SAME**

[01] FIELD OF THE INVENTION

[02] The invention generally relates to methods for fabricating integrated circuits (ICs) and semiconductor devices and the resulting structures. More particularly, the invention relates to metal oxide semiconductor field effect transistor (MOSFET) devices and methods for making such devices. Even more particularly, the invention relates to self-aligned trench MOSFET devices and methods for making such devices.

[03] BACKGROUND OF THE INVENTION

[04] There are numerous devices comprising silicon layers containing deep or high aspect ratio trenches. Forming such deep trenches in the silicon layer of these devices provides many novel and promising structures. The types of devices containing such trenches include the numerous types of silicon-based MEMS devices, as well ICs.

[05] In IC fabrication, devices such as transistors may be formed on a semiconductor wafer or substrate, which is typically made of silicon. MOSFET devices are widely used in numerous applications, including automotive electronics, disk drives and power supplies. Generally, these devices function as switches, and they are used to connect a power supply to a load. It is important that the resistance of the device be as low as possible when the switch is closed. Otherwise, power is wasted and excessive heat may be generated.

[06] One type of MOSFET is described in U.S. Patent No 6,084,268, the disclosure of which is incorporated herein by reference. In another type of MOSFET, the gate is formed in a trench. See, for example, U.S. Patents No. 6,084,264, 5,998,833, 5,998,836, 5,998,837, 6,049,108, 6,051,488, 5,895,952, 6,204,533, and 6,090,716, the disclosures of which are incorporated herein by reference. An example of such a device is illustrated in Figures 1 and 2, with Figure 1 depicting a cross-sectional view of a single cell of a MOSFET 100, and Figure 2 depicting a plan view of the cell. Gates 102 and 104 are formed in trenches and surrounded by gate oxide layers 106 and 108, respectively. The trench gate is often formed in a grid pattern including an array of polygonal shapes (one section of which is shown in Figure 2), the grid representing a single interconnected gate (gates 102 and 104 being the same). The trench gate may also be formed as a series of distinct parallel stripes.

[07] MOSFET 100 is formed in an N-epitaxial layer 110. A N+ source region 112 is formed at the surface of epitaxial layer 110. A P+ contact region 114 is also formed at the surface of epitaxial layer 110. A P-body region 116 is located below N+ source region 112 and P+ contact region 114. A metal source contact 118 contacts the source region 112 and shorts the source region 112 to the P+ contact region 114 and P body region 116.

[08] The N-epitaxial layer 110 is formed on a substrate 120, and a drain contact (not shown) is located at the bottom of the substrate 120. The contact for the gates 102 and 104 is likewise not shown, but it is generally made by extending the conductive gate

material outside of the trench and forming a metal contact at a location remote from the individual cells. The gate is typically made of phosphorus or boron doped polysilicon.

[09] A region 111 of N-epitaxial layer 110 between the substrate 120 and the P body 116 is generally more lightly doped with N-type impurities than substrate 120. This increases the ability of MOSFET 100 to withstand high voltages. Region 111 is sometimes referred to as a "lightly doped" or "drift" region ("drift" referring to the movement of carriers in an electric field). Drift region 111 and substrate 120 constitute the drain of MOSFET 100.

[10] One feature making the trench configuration attractive is that the current flows vertically through the channel of the MOSFET. This permits a higher packing density than MOSFETs such as described in U.S. Patent No 6,084,268, where the current flows horizontally through the channel and then vertically through the drain. Greater cell density generally means more MOSFETs can be manufactured per unit area of the substrate and, therefore, a lower on-resistance ($R_{DS(on)}$) for the device. Increasing the cell density also allows the individual transistors to be closer together, requiring less area on a substrate or wafer and reducing the cost of manufacturing the device.

[11] One problem standing in the way of increasing the cell density by manufacturing the transistors closer together is the alignment tolerance. The alignment tolerance is the amount of over-sizing needed to compensate for variations in the alignment. For example, as illustrated in Figure 3, after the gate conductor 10 is placed in the trench 11, an isolation dielectric 4 is placed over the conductor 10 to electrically isolate the gate conductor from other conductive layers that will be placed on the substrate 12 during

subsequent processing. Because the isolation dielectric 4 is only over the gate—and other selected—portions of the device, a masking and etching step using patterned photoresist mask 22 is employed to define the portions of the device where the isolation dielectric 4 need not be present, such as upper surface 13 of substrate 12. To compensate for the variations in the alignment of the trenches, an excess amount of the etched isolation dielectric 4A remains for adequate isolation, as depicted in Figure 4. Because of this alignment tolerance 6, the transistors cannot be placed closer together and the cell density cannot be increased to yield the advantages mentioned above.

[12]

SUMMARY OF THE INVENTION

[13]

The present invention provides self-aligned trench MOSFETs and methods for manufacturing the same. By having a self-aligned structure, the number of MOSFETs per unit area—the cell density—is increased, increasing performance and making the MOSFETs cheaper to produce. The self-aligned structure for the MOSFET is provided by making the sidewall of the overlying isolation dielectric layer substantially aligned with the sidewall of the gate conductor. Such an alignment can be made through any number of methods such as using a dual dielectric process, using a selective dielectric oxidation process, using a selective dielectric deposition process, or a spin-on-glass dielectric process.

[14]

[15]

[16]

[17]

[18]

[19]

[20]

[21]

[22]

[23]

[24]

[25]

that the present invention can be practiced without employing these specific details. Indeed, the present invention can be practiced by modifying the illustrated structure and method, and can be used in conjunction with apparatus and techniques conventionally used in the industry. For example, while the invention is described with reference to MOSFET devices, it could be modified for other devices formed in trenches which need to be isolated, such as bipolar devices, BDCMOS devices, or other types of transistor structures. As well, although the devices of the invention are described with reference to a particular type of conductivity (P or N), the devices can be configured with the opposite type of conductivity (N or P, respectively) by appropriate modifications.

[26] Figures 5-14 illustrate one aspect of the process for fabricating a MOSFET in accordance with this invention. Other processes resulting in substantially similar structures can, of course, be used in the invention. As shown in Figure 5, the process begins with a substrate 12. Any substrate known in the art can be used in the invention. Suitable substrates include silicon wafers, epitaxial Si layers, polysilicon layers, bonded wafers such as used in silicon-on-insulator (SOI) technologies, and/or amorphous silicon layers, all of which may be doped or undoped.

[27] In one aspect of the invention, the substrate 12 in the invention is a single crystal silicon wafer 12A having at least one epitaxial ("epi") Si layer 15 located in an upper region thereof. If desired, more than one epitaxial layer can be provided on the upper surface. The epitaxial layer(s) can be provided using any known process in the art, including any known epitaxial deposition process. The epitaxial layer(s) can remained

undoped or can be doped with a desired dopant to the desired concentration using any suitable process known in the art.

[28] Then, as illustrated in Figure 5, a termination junction 20 is formed in an upper region of the epitaxial layer 15. Although any method known in the art can be used, the following process is used in the invention. A photoresist layer is deposited and then patterned as known in the art to form a patterned resist layer 23. The patterned resist layer 23 is then used as an implant mask during the implantation process used to create the termination junction. The implantation process forms a termination implant region 21 by implanting a suitable dopant to obtain the desired profile and concentration. Any dopant known in the art can be employed in the invention depending on the desired type of conductivity, e.g., P or N. Examples of suitable dopants include such as arsenic, boron, phosphorus, aluminum, antimony, beryllium, gallium, germanium, gold, magnesium, tellurium, tin, or a combination of these dopants. In one aspect of the invention, boron is implanted at an energy of about 30 to about 60 KeV with a dopant concentration of about 1×10^{14} to about 5×10^{15} . The resist layer is then stripped using conventional processing, like a wet etch process using hot sulfuric acid and hydrogen peroxide.

[29] Next, as illustrated in Figure 6, an oxide layer 24 is provided over the entire surface of the substrate 12. While any known process can be employed in the invention, the oxide layer 24 is preferably deposited by heating the substrate 12 in an oxidizing ambient such as oxygen or steam at 900-1100°C. Then, the dopant in the termination implant region 21 is driven into the substrate to form termination junction 20 to the

desired depth, e.g., a depth of about 2,000 to about 10,000 angstroms. While any process obtaining this desired dopant profile can be used, the dopant is preferably driven by heat treatment in a furnace at a temperature of about 1000 to about 1100 degrees Celsius.

[30] Next, as illustrated in Figure 7, a field oxide 25 is provided from oxide layer 24. The field oxide 25 is patterned and etched from the active areas of the device using any known methods in the art. For example, a photoresist layer can be deposited and patterned. The resulting resist mask 26 is then used during an etching process to remove undesired portions of the oxide layer 24. The resist layer is then stripped using conventional processing, like a wet etch process using hot sulfuric acid and hydrogen peroxide.

[31] Next, as depicted in Figures 8 and 9, the trenches 11 in substrate 12 are made by a masking and etching process. Mask 24 can comprise either a single layer or multiple layers as known in the art. See, for example, U.S. Patent Nos. 6,103,635 and 6,121,154, the disclosures of which are incorporated herein by reference. In one aspect of the invention, mask 24 can comprise, for example, any suitable photoresist material. The thickness and other characteristics of the photoresist are well known and can be optimized as known in the art.

[32] As depicted in Figure 8, mask 24 contains a pattern of openings 31 that expose the upper surface 13 of substrate 12. The openings 31 generally correspond to the locations where the trenches 11 will be formed. To form the openings, the mask 24 is patterned during a photolithographic process as known in the art. The patterning process removes portions of the photoresist material in the location of openings 31.

[33] With openings 31 formed in mask 24, trenches 11 are then formed in substrate 12 using any suitable process known in the art. In one aspect of the invention, the trenches are formed via a chlorine-based etching process. The etching process begins by forming upper sidewall portions in the epitaxial silicon layer 15. The process then continues to produce lower sidewall portions and rounded bottom corners. In one aspect of the invention, the depth of the trenches can range from about 0.5 to about 3 microns. The parameters of the etching process are controlled to preferably form round bottom corners, smooth surfaces, and flat and clean trench bottom surfaces, thereby maintaining the integrity of the MOSFET device characteristics using the trenches.

[34] After forming the trenches, the mask 24 is removed by any suitable process known in the art, resulting in the structure illustrated in Figure 9. The photoresist material can be removed using a conventional stripping solution such as $\text{H}_2\text{SO}_4/\text{H}_2\text{O}_2$ or an O_2 ash. Following removal of resist mask 24, the epitaxial silicon layer 15 can be dipped in a diluted acidic solution to remove any polymer and/or passivation present in the trenches. Further cleaning can be performed as known in the art.

[35] As shown in Figure 10, an oxide layer 27 is then provided on the top surface of the structure, including the trench. This oxide will be used as gate oxide 18 and therefore must be of a high quality. Any suitable method known in the art—including oxidation and deposition—yielding a high quality oxide layer can be used to provide oxide layer 27. In one aspect of the invention, oxide layer 27 is provided by oxidation or deposition. The thickness of oxide layer 27 can range from about 30 to about 1000 angstroms.

[36] Next, a conductive layer 28 is deposited to fill and overflow the remaining portions of the trench 11. This conductive layer 28 can be suitable material that can be used as a gate conductor 10, such as metal or polysilicon. In one aspect of the invention, the conductive layer is preferably polysilicon. The conductive layer 28 can be deposited using any known deposition process, including chemical vapor deposition process. Optionally, the conductive layer 28 can be doped with any suitable dopant to the desired concentration. In one aspect of the invention, such as when the conductive layer is polysilicon, the conductor layer is doped with P or B to a concentration of about 5×10^{19} by ion implantation or thermal processing.

[37] Excess (and unneeded) portions of the conductive layer 28 are then removed using a photoresist mask 29 to form gate conductor 10. In one aspect of the invention, enough of the conductive layer 28 is removed so the upper surface 30 of the resulting gate conductor 10 is lower than the upper surface 13 of the substrate 12. The distance between the upper surface 13 of the substrate 12 and the upper surface 30 of the gate conductor 10 can range from about 0.25 to about 0.8 microns, and in one aspect of the invention, is preferably about 0.5 microns. The excess portions of the conductive layer 28 can be removed through any process known in the art, such as a suitable etching process. The portions of the oxide layer 27 outside the trenches 11 can then be removed through any process known in the art, such as a suitable etching process. Then, the photoresist mask 29 is stripped as known in the art.

[38] Next, as illustrated in Figures 11-14, implant region 32 and channel regions 33 are created in an upper region of substrate 12. In one aspect of the invention, any suitable

doping process that will obtain the desired profile and concentration can be used to create the implant and channel regions. For example, such regions can be created by implanting P, As, or B at about 30 to about 80 KeV with a dose of about 3×10^{13} . If necessary, a mask can be used to limit the doping process to specific areas of the substrate. Then, the dopant is driven into the substrate by heating for about 20 to about 120 minutes at a temperature of about 1100 degrees Celsius.

[39] Next, source region 34 is created in an upper region of substrate 12 near gates 16 and 17. In one aspect of the invention, the source region 34 can be created using any suitable doping process that will obtain the desired profile and concentration. For example, the source region can be created by implanting As, P, or B at about 80KeV with a dose of about 1×10^{16} . A source mask 41 made of a suitable photoresist material is used to limit the doping process to those areas of the substrate where the source region will be created. After the doping process to create the source regions, the source mask is removed using any suitable stripping process. Then, the dopant is driven into the substrate by heating for about 60 minutes at a temperature of about 800 to about 1000 degrees Celsius.

[40] Next, body contact region 35 is created in an upper region of substrate 12. In one aspect of the invention, the body contact region 35 can be created using any suitable doping process that will obtain the desired profile and concentration. For example, the body contact region can be created by implanting B, P, or As at about 40-100KeV with a dose of about 1×10^{14} to about 8×10^{15} . Then, the dopant is driven into the substrate by a standard furnace anneal or a rapid thermal anneal.

[41] After the above steps have been performed, a self-aligned dielectric layer (or isolation cap 38) is provided over gate conductor 10. In one aspect of the invention, the isolation cap is created in the following manner, the "dual dielectric process." In this aspect of the invention, as shown in Figures 15 and 16, a first dielectric layer 36 is provided over the surface of the substrate 12. The first dielectric layer can be provided using any suitable process in the art, such as an oxidation or deposition process like PECVD. The first dielectric layer 36 can comprise any dielectric material such as PSG, BPSG, or an undoped LTO (low temperature oxide) layer. Preferably, the first dielectric layer is PSG.

[42] The first dielectric layer 36 is doped so that it etches faster than the second dielectric layer 37 (described below). The first dielectric layer 36 can be doped after it is deposited or while it is being deposited (*in situ*). Preferably, when first dielectric layer 36 is made of PSG, it is doped with P to a concentration of about 1 to about 3%.

[43] Next, a second dielectric layer 37 is provided over the surface of the first dielectric layer 36. The second dielectric layer 37 can be provided using any suitable process in the art such as a deposition process like PECVD. The second dielectric layer 37 can comprise any dielectric material such as PSG or BPSG. Preferably, the second dielectric layer is BPSG. The second dielectric layer (and accompanying dopant as described below) is also selected so that it reflows more readily than the first dielectric layer 36.

[44] The second dielectric layer 37 is doped so that it etches slower than the first dielectric layer 36. The second dielectric layer can be doped after it is deposited or while it is being deposited (*in situ*). Preferably, when second dielectric layer 37 is made of BPSG, it is doped with B to a concentration of about 3-5% and P to a concentration of about 3-5%.

[45] After both the first and second dielectric layers have been deposited, both dielectric layers are flowed. The flow process will create a planar upper surface for the dielectric layers, as depicted in Figure 15. The flow process selected can be any known in the art that will create the substantially planar upper surface, like heating for 100 minutes at a temperature of about 850 to about 900 degrees Celsius.

[46] Next, the first and second dielectric layers are removed to leave an isolation cap 38 in the trench on the gate conductor 10. The dielectric layers can be removed using any suitable process, such as an isotropic etching process using BOE or anisotropic etching processes such as a dry plasma etch. The etching process removes the dielectric layers from the regions where they need not be present, as well as leaves a substantially planar upper surface of isolation cap 38.

[47] In a variation on this aspect of the invention, a single dielectric layer can be used instead of two dielectric layers. The single dielectric layer can be doped or undoped. The single dielectric layer is then flowed and isotropically etched in vertical direction until the upper surface of the substrate is exposed. The process in this variation leaves an isolation cap in the trench over the gate conductor because the thickness of the reflowed single

dielectric layer is greater in the trench than outside the trench and the isotropic etch will remove the same amount of dielectric layer regardless of location.

[48] In another aspect of the invention, the isolation cap 38 is provided in an alternative manner, the "selective dielectric oxidation process." In this aspect of the invention, as shown in Figure 17, the mask 24 used while etching the trenches in the substrate 12 is made of nitride. After forming the trenches 11, gate oxide 18, and gate conductor 10, a single dielectric layer is grown over the gate conductor 10. Preferably, the dielectric layer 39 is grown by oxidation. For example, this dielectric layer can be grown by oxidation in oxygen and hydrogen in a high-temperature oxidation furnace at a temperature of about 900 to about 1100 degrees Celsius for about 30 to about 90 minutes.

[49] Because of the presence of mask 24 containing a nitride, the oxidation process does not oxidize portions of the substrate upper surface 13 other than in the trenches 11. Thus, after the oxidation process, the mask 24 containing the nitride can be removed by any suitable process, such as a wet etch using hot phosphoric acid. With mask 24 removed, isolation cap remains over gate conductor 10.

[50] In another aspect of the invention, the isolation cap 38 is provided in yet another manner, the "selective dielectric deposition process." In this aspect of the invention, the mask 24 used while etching the trenches in the substrate 12 is again made of nitride. After forming the trenches 11, gate oxide 18, and gate conductor 10, a single dielectric layer is grown over the gate conductor 10. Preferably, the dielectric layer is selectively deposited by any suitable deposition process known in the art. For example, the dielectric layer can be deposited by SACVD and then flowing TEOS and ozone in a temperature of

about 400 to about 500 degrees Celsius and a pressure ranging from about 200 to about 600 torr, thereby depositing a non-organic silicon dioxide layer containing only residual amounts of carbon atoms.

[51] Because the deposition is a selective deposition process, the single dielectric layer is only deposited over gate conductor 10. Thus, after the deposition process, the mask 24 containing the nitride can be removed by any suitable process, such as a wet etch using hot phosphoric acid. With mask 24 removed, isolation cap remains over gate conductor 10.

[52] In another aspect of the invention, the isolation cap 38 is provided in still yet another manner, the "SOG dielectric process." In this aspect of the invention, as illustrated in Figure 18, after the gate conductor 10 is formed by removing the desired portions of the conductive layer 28, an organic dielectric material such as spin-on-glass (SOG) layer is applied to the substrate as known in the art. See, for example, U.S. Patent No. 6,090,716, the disclosure of which is incorporated herein by reference. The SOG layer is applied in such a manner so that the SOG remains in the trenches 11 to form the isolation cap 38. As well, the SOG can be applied so that it optionally remains on the portion of conductive layer 28 overlying the field oxide.

[53] After the above processes are concluded, conventional processing can continue to finish the MOSFET devices. As well, other processing can continue to finish other processing needed to complete other parts of the semiconductor device.

[54] Having described the preferred embodiments of the present invention, it is understood that the invention defined by the appended claims is not to be limited by

particular details set forth in the above description, as many apparent variations thereof are possible without departing from the spirit or scope thereof.

209110-40013.003